**ASSIGNMENT 15.3**

**1.Explain the working of SQOOP in brief.**

Sqoop is a tool designed to transfer data between Hadoop and relational database servers. It is used to import data from relational databases such as MySQL, Oracle to Hadoop HDFS, and export from Hadoop file system to relational databases.

**Sqoop:** “SQL to Hadoop and Hadoop to SQL”

The import tool imports individual tables from RDBMS to HDFS. Each row in a table is treated as a record in HDFS. All records are stored as text data in text files or as binary data in Avro and Sequence files.

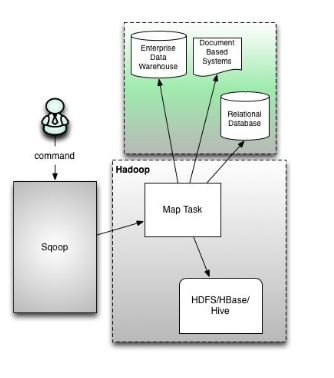
The export tool exports a set of files from HDFS back to an RDBMS. The files given as input to Sqoop contain records, which are called as rows in table. Those are read and parsed into a set of records and delimited with user-specified delimiter.

Sqoop automates most of the process, depends on the database to describe the schema of the data to be imported. Sqoop uses MapReduce framework to import and export the data, which provides parallel mechanism as well as fault tolerance. Sqoop makes developers life easy by providing command line interface. Developers just need to provide basic information like source, destination and database authentication details in the sqoop command. Sqoop takes care of remaining part.

Sqoop provides many salient features like:

1. Full Load
2. Incremental Load
3. Parallel import/export
4. Import results of SQL query
5. Compression
6. Connectors for all major RDBMS Databases
7. Kerberos Security Integration
8. Load data directly into Hive/Hbase
9. Support for Accumulo

Sqoop is Robust, has great community support and contributions. Sqoop is widely used in most of the Big Data companies to transfer data between relational databases and Hadoop.

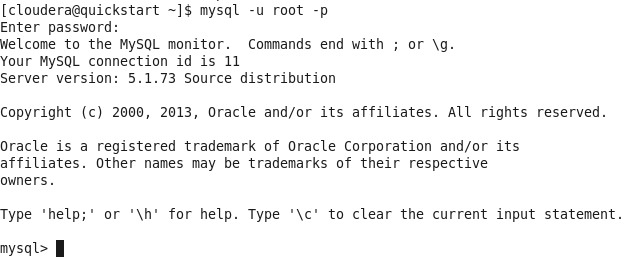


**2. Import a table from Mysql to HBase table using Sqoop**

#### Login to MySQL shell:

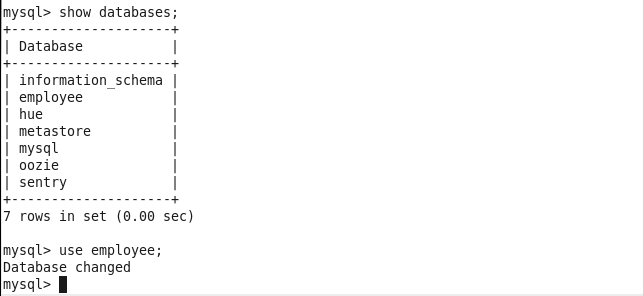
|  |  |
| --- | --- |
| 1 | mysql -u root –p |

**Password: cloudera**



#### Show databases:

|  |  |
| --- | --- |
| 1 | show databases; |
| 1 | Use employee; |



Describe table:

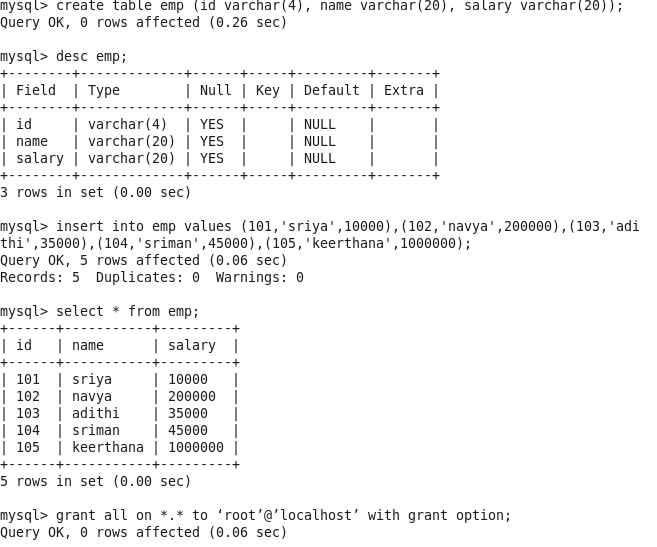
|  |  |
| --- | --- |
| 1 | Desc emp; |

#### Display the contents of the table employee:

|  |  |
| --- | --- |
| 1 | select \* from emp; |

 Grant all permission:

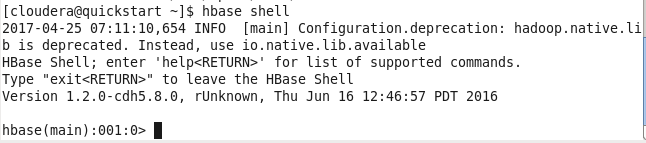
|  |  |
| --- | --- |
| 1 | grant all on \*.\* to ‘root’@’localhost’ with grant option; |



#### Enter Hbase shell:

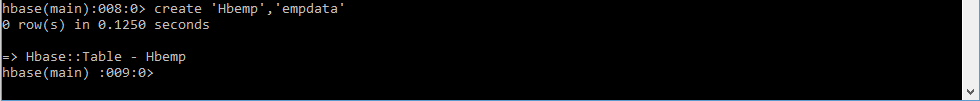
Use below command to enter HBase shell.

|  |  |
| --- | --- |
| 1 | hbase shell |



#### Create table:

|  |  |
| --- | --- |
| 1 | create ‘Hbemp’,’empdata’ |



#### Sqoop import command:

|  |  |
| --- | --- |
| 1  2 | sqoop import --connect jdbc:mysql://localhost/employee --username root --password cloudera --table emp  --hbase-table Hbemp --column-family empdata --hbase-row-key id -m 1 |

#### Scan HBase table:

|  |  |
| --- | --- |
| 1 | scan 'Hbemp' |

